

Online MSU DATABASE Helps World Research

By Bonnie Coblentz

When scientists around the world need information on gene function in agriculturally important crops, they turn to an online database developed at Mississippi State University.

Known as AgBase, this database catalogs the functions of genes in the genome sequences of plants and animals. Dr. Shane Burgess, a College of Veterinary Medicine researcher, was the lead collaborator on the project.

Burgess said researchers around the world are sequencing genomes and coming up essentially with lists as part that make up life.

"It doesn't help us much unless we understand what all those parts are supposed to do, how they work and how they work together in defined contexts," Burgess said.

The next step is to understand life at this complicated genomic level. Technology today allows researchers to develop massive amounts of data quantifying how the genes in the genome are all expressed under given conditions. This work is called "functional genomics."

In biology, the component parts are proteins, the building blocks for molecular machines inside the cells that make up an organism. Protein machines are the basis of all life functions.

"When genomes are sequenced and data sets are generated, all we end up with is the equivalent of dirty laundry lists," Burgess said. "We want to go from the dirty laundry lists to be able to infer what was going on at the party the night before."

To do that, each piece of information generated about gene function must be defined in a standardized fashion, the possible variations noted and information included about where in the cell these proteins are found.

"We define all these functions, as well as where they occur in the cell, using a dictionary of defined terms called the gene ontology," Burgess said. "Some very clever people studying mice, fruit flies and yeast first came up with this ontology for describing life. Now it is a global effort that defines life in bacteria, viruses, animals, plants, in fact everything that is alive."

Once the gene ontology terms are defined, they are given a digital code so computers can be used to help analyze the huge functional genomics data sets that define life.

AgBase is the first defined, online, mapped database for this particular genome functional information in agriculture. Burgess said it is the internationally recognized tool for modeling



Shane Burgess and Susan Bridges collaborate on AgBase. (Photo by Tom Thompson)

functional genomics data in agricultural species.

"We are the go-to place for all of agriculture for this information and computational tools that allow you to model life based on functional genomics data," Burgess said.

Burgess said the project began as an idea in November 2004, and AgBase was launched in June 2005.

"This project would not have been possible without the excellent and multidisciplinary collaborations that we have at MSU," Burgess said.

Burgess collaborated on AgBase with Susan Bridges, a professor in the Department of Computer Science and Engineering, and Dawn Luthe in biochemistry. Dr. Fiona McCarthy at the veterinary college and Nan Wang in computer sciences also are researchers in this project.

Bridges is responsible for the computer support of AgBase, including building the database itself and the web interface that allows researchers worldwide to use it.

"We were in the right place at the right time with the right combination of people," Bridges said.

The long-term goal of AgBase is to serve the post-genome biology needs of the agricultural community and for biomedical and other researchers primarily using traditionally agricultural species as biomedical models.

AgBase is recognized by the international Gene Ontology Consortium as the official source of this data. Recently, AgBase has begun to be a central player for determining the names of genes in agricultural species worldwide.